

Current density operator in systems with non-parabolic, position-dependent energy bands

William R. Frensley

*Electrical Engineering, University of Texas at Dallas, Richardson, TX 75080**

Abstract

The present manuscript was written in 1994 and was not published. It addresses the form that the quantum-mechanical current density must take in mesoscopic treatments of semiconductor heterostructures, in which the electron dispersion relations are non-parabolic and position dependent, rendering the textbook expressions inapplicable. The approach is to derive the continuity equation for the specific model under consideration, using generalizations of Green's identities to higher-order derivatives and to discrete models of different topological structure. A new addendum addresses two issues of more current interest: the use of irregular meshes in discrete formulations, and the identification of the Heisenberg velocity operator to evaluate current density. It is demonstrated that on discrete domains the velocity operator fails to satisfy a sensible continuity equation, and therefore cannot be identified with the current density.

PACS numbers: 73.40.-c, 71.25.-s

* frensley@utdallas.edu

ORIGINAL 1994 ABSTRACT

In semiconductor heterostructures the electron dispersion relations are non-parabolic and vary with position. The non-parabolicity is described by effective Hamiltonians which either include higher-order derivatives or couple several basis states. As a result, the current density operator is not simply related to the gradient. By generalizing Green's identity to higher-order derivatives and to difference relations, the appropriate form of this operator is derived for all of the commonly-used band structure representations.

I. INTRODUCTION

The wave mechanics of semiconductor heterostructures is complicated by the fact that the electron dispersion relation (or energy band structure) is generally non-parabolic at modest energies and necessarily varies with position. Under such circumstances, the form of the current density operator is no longer simply a symmetrized gradient.

The form of the particle current density operator J is clearly constrained by the group-velocity theorem [1], so that the expectation value of J on a state of definite wavevector k is

$$\langle k|J|k\rangle = v_g\langle k|k\rangle = \partial E/\partial k\langle k|k\rangle/\hbar. \quad (1)$$

This equation, together with the band energies and eigenstates, in principle determines the form of J . It is, however, much more convenient to directly derive J from the Hamiltonian for a given problem. One does so by evaluating the time derivative of the probability density:

$$\partial\psi^*\psi/\partial t = (1/i\hbar)[\psi^*(H\psi) - (H\psi^*)\psi]. \quad (2)$$

Green's identity [$f\nabla^2 g - g\nabla^2 f = \nabla \cdot (f\nabla g - g\nabla f)$], or a generalization thereof, is then invoked to write the right-hand side of (2) as the divergence of the current density.

Heterostructures are most often described at a "mesoscopic" level where the microscopic (smaller than the atomic diameter) behavior of the wavefunction can be factored out. In order to realistically describe the non-parabolic dispersion relation, the resulting effective Hamiltonian must be more elaborate than a simple Laplace operator, and Green's identity must be correspondingly be generalized to derive the form of J . The commonly-used mesoscopic models can be classified as effective-mass or tight-binding approaches.

II. THE CURRENT DENSITY IN SPECIFIC REPRESENTATIONS

For the purposes of the present discussion, we will assume that we possess a hermitian effective Hamiltonian which is valid throughout the structure, including any interval containing an abrupt heterointerface. The issues thus presumed to have been resolved have traditionally been posed as the definition of matching conditions for the mesoscopic wavefunction. These matching conditions are frequently “derived” from the continuity of J [2–4], but this condition is not sufficient to uniquely determine them [5]. The Hamiltonian and the matching conditions are equivalent pieces of information, in the sense that either one may be derived from the other. In contrast, the continuity of J follows solely from the hermiticity of H [6].

For the sake of simplicity, only the current density in one dimension will be considered. Extension to three-dimensional structures is in all cases straightforward, if somewhat tedious.

A. Wannier-Slater Effective Mass Theory

In the approach to effective-mass theory proposed by Wannier [7] and expounded by Slater [8], the microscopic wavefunction ψ is expanded as a linear combination of localized Wannier functions, each of which is centered within a different unit cell. The expansion coefficients Ψ_i can be regarded as values of a discrete lattice function which is interpolated between lattice points by a continuous function $\Psi(z)$, for which an effective-mass Schroedinger equation is derived. If the electron dispersion relation can be expanded as

$$E(k) = \sum_{n=0}^N A_n(z) k^{2n}, \quad (3)$$

then the effective Hamiltonian is

$$H^{(\text{WS})}\Psi = \sum_{n=0}^N (-1)^n \frac{\partial^n}{\partial z^n} A_n(z) \frac{\partial^n \Psi}{\partial z^n} + V(z)\Psi. \quad (4)$$

To derive the appropriate J for this Hamiltonian, we require a generalized Green’s identity:

$$f \frac{\partial^n}{\partial z^n} A \frac{\partial^n g}{\partial z^n} - g \frac{\partial^n}{\partial z^n} A \frac{\partial^n f}{\partial z^n} = \frac{\partial}{\partial z} \sum_{j=0}^{n-1} (-1)^j \left[\frac{\partial^j f}{\partial z^j} \frac{\partial^{n-j-1}}{\partial z^{n-j-1}} A \frac{\partial^n g}{\partial z^n} - \frac{\partial^j g}{\partial z^j} \frac{\partial^{n-j-1}}{\partial z^{n-j-1}} A \frac{\partial^n f}{\partial z^n} \right]. \quad (5)$$

(This identity is readily proven by expanding the derivative on the right-hand side; the summation then becomes a telescoping series.) The value of current density is thus

$$\langle J_z^{(\text{WS})} \rangle(z) = \frac{i}{\hbar} \sum_{n=1}^N \sum_{j=0}^{n-1} (-1)^{n-j} \left[\frac{\partial^j \Psi^*}{\partial z^j} \frac{\partial^{n-j-1}}{\partial z^{n-j-1}} A_n \frac{\partial^n \Psi}{\partial z^n} - \frac{\partial^j \Psi}{\partial z^j} \frac{\partial^{n-j-1}}{\partial z^{n-j-1}} A_n \frac{\partial^n \Psi^*}{\partial z^n} \right]. \quad (6)$$

Some formulations of the matching conditions produce terms in the Hamiltonian of the form

$$H^{(2n+1)} = \frac{i}{2} (B_{2n+1} \partial^{2n+1} / \partial z^{2n+1} + \partial^{2n+1} / \partial z^{2n+1} B_{2n+1}). \quad (7)$$

Their contribution to the current density may be readily derived from another identity:

$$f \frac{\partial^{2n+1} g}{\partial z^{2n+1}} + g \frac{\partial^{2n+1} f}{\partial z^{2n+1}} = \frac{\partial}{\partial z} \sum_{j=0}^{2n} (-1)^j \frac{\partial^j f}{\partial z^j} \frac{\partial^{2n-j} g}{\partial z^{2n-j}}, \quad (8)$$

leading to contributions to the current density of

$$\langle J^{(2n+1)} \rangle(z) = -\frac{1}{2\hbar} \sum_{j=0}^{2n} (-1)^j \left[\frac{\partial^j B_{2n+1} \psi^*}{\partial z^j} \frac{\partial^{2n-j} \psi}{\partial z^{2n-j}} + \frac{\partial^j \psi^*}{\partial z^j} \frac{\partial^{2n-j} B_{2n+1} \psi}{\partial z^{2n-j}} \right]. \quad (9)$$

It appears (based upon an examination of some low-order cases) that any apparently hermitian differential operator [9] can be manipulated into an expression containing only terms of the form (4) and (7).

B. Luttinger-Kohn Effective Mass Theory

The Luttinger-Kohn approach to effective-mass theory [10–12] more conveniently includes the effects of several bands. In this scheme the microscopic wavefunction is decomposed into the $\mathbf{k} = 0$ Bloch functions and a set of slowly-varying envelope functions $\chi_m(z)$, m being a band index. Differing numbers of bands may be included, with perhaps the most general scheme being that derived by Bastard [11]. By regrouping the various terms of Bastard's Hamiltonian with respect to the derivatives, it can be written in the form [13] (summations implied over repeated indices)

$$(H^{(\text{LK})} \chi)_l = \left[-\frac{\partial}{\partial z} A_{lm} \frac{\partial}{\partial z} - \frac{i}{2} (B_{lm} \frac{\partial}{\partial z} + \frac{\partial}{\partial z} B_{lm}^\dagger) + C_{lm} \right] \chi_m, \quad (10)$$

where A and C are hermitian matrices and B a general matrix. A , B , and C are indexed by the band label m , and are z -dependent in a heterostructure. Using the above identities, the current density is readily shown to be

$$\langle J_z^{(\text{LK})} \rangle(z) = \frac{1}{i\hbar} \left[\chi_l^* A_{lm} \frac{\partial \chi_m}{\partial z} - \frac{\partial \chi_l^*}{\partial z} A_{lm} \chi_m + \frac{i}{2} \chi_l^* (B + B^\dagger)_{lm} \chi_m \right]. \quad (11)$$

Similar expressions have been obtained by Altarelli [14] and by Burt [12].

C. Tight-Binding Theories

In the tight-binding approach [15], the wavefunction is expanded in terms of a set of localized states $m = 1, \dots, M$ in each atomic layer j

$$|\psi\rangle = \sum_{j,m} c_{jm} |jm\rangle. \quad (12)$$

The coefficients c_{jm} can be thought of as forming a block-structured vector \mathbf{c} with vector elements $\mathbf{c}_j = [c_{j1}, \dots, c_{jM}]^T$. The Hamiltonian then becomes a block-structured matrix $\mathbf{H}^{(\text{TB})}$, of which the diagonal blocks $\mathbf{H}_{ii}^{(\text{TB})}$ are hermitian and describe interactions within a plane and the off-diagonal blocks $\mathbf{H}_{ij}^{(\text{TB})}$ are not necessarily hermitian ($\mathbf{H}_{ij}^{(\text{TB})} = \mathbf{H}_{ji}^{(\text{TB})\dagger}$) and describe the coupling between planes. If only nearest-neighbor interactions are included, $\mathbf{H}^{(\text{TB})}$ is block-tridiagonal [16, 17] ($\mathbf{H}_{ij}^{(\text{TB})} \neq 0$ only for $j = i - 1, i, i + 1$).

Because the tight-binding representation is intrinsically discrete, we need to modify somewhat our concepts of probability and current density. A total probability density ρ_i is associated with each atomic plane i , and is equal to $\mathbf{c}_i^\dagger \mathbf{c}_i$. The current density represents the flux between adjacent planes; we will write the flux between planes i and $i + 1$ as $J_{i+1/2}$ [18]. Applying (2) to the tight-binding Hamiltonian $\mathbf{H}^{(\text{TB})}$ and assuming only nearest-neighbor interactions, we get

$$\frac{\partial}{\partial t} \mathbf{c}_i^\dagger \mathbf{c}_i = \frac{1}{i\hbar} \left[\mathbf{c}_i^\dagger \mathbf{H}_{i,i-1}^{(\text{TB})} \mathbf{c}_{i-1} + \mathbf{c}_i^\dagger \mathbf{H}_{i,i+1}^{(\text{TB})} \mathbf{c}_{i+1} - \mathbf{c}_{i-1}^\dagger \mathbf{H}_{i-1,i}^{(\text{TB})} \mathbf{c}_i - \mathbf{c}_{i+1}^\dagger \mathbf{H}_{i+1,i}^{(\text{TB})} \mathbf{c}_i \right]. \quad (13)$$

This can be written as a discrete continuity equation,

$$\partial \rho_i^{(\text{TB})} / \partial t = \langle J^{(\text{TB})} \rangle_{i-1/2} - \langle J^{(\text{TB})} \rangle_{i+1/2}, \quad (14)$$

if $\langle J^{(\text{TB})} \rangle$ is identified as

$$\langle J^{(\text{TB})} \rangle_{i+1/2} = \frac{1}{i\hbar} \left(\mathbf{c}_{i+1}^\dagger \mathbf{H}_{i+1,i}^{(\text{TB})} \mathbf{c}_i - \mathbf{c}_i^\dagger \mathbf{H}_{i,i+1}^{(\text{TB})} \mathbf{c}_{i+1} \right). \quad (15)$$

Here we see the machinery of Green's identity operating in a discrete space.

A variation of the tight-binding scheme is the ‘‘Wannier Orbital Model’’ [19], which draws upon the discrete form of the Wannier-Slater theory. Interactions with remote neighbors are included to fit the $E(k)$ dispersion, and typically $M = 1$, so that only one band is modeled. Inserting such a Hamiltonian into (2) leads to many terms, which cannot be associated with a particular position. Thus, the notion of a local current density disappears, due to the

direct interactions between remote sites. Instead, we may define an antisymmetric current matrix with elements

$$\langle J^{(\text{WO})} \rangle_{ij} = (i/\hbar) \left(c_i^* H_{ij}^{(\text{WO})} c_j - c_j^* H_{ji}^{(\text{WO})} c_i \right), \quad (16)$$

where $\langle J \rangle_{ij}$ is the current flowing out of site i into site j . The nonlocal continuity equation is then

$$\partial \rho_i^{(\text{WO})} / \partial t = - \sum_j \langle J^{(\text{WO})} \rangle_{ij}. \quad (17)$$

III. SUMMARY

The form of the current density has been derived for all of the usual models of heterostructure electronic states. While one can usually find a way to obtain correct answers in a manual calculation without knowledge of the general expressions for J presented here, they provide a useful check on the results. The use of these expressions becomes more necessary if one seeks to develop the machinery for automatic computations (numerical or symbolic) applicable to a wide variety of heterostructures.

The reader will have noticed that, in conformance to the practice in quantum-mechanics texts, expressions for the expectation values of J have been presented, not expressions for the operator itself. It is very difficult to represent J as an ordinary quantum-mechanical operator. Actually, J is much more naturally expressed as a superoperator which acts upon a density operator [20]. All of the results presented here may be derived much more elegantly in a superoperator formalism, at the cost of an unfamiliar notation.

REFERENCES

- [1] H. Kroemer, Proc. IEEE **63**, 988 (1975).
- [2] W. A. Harrison, Phys. Rev. **123**, 85 (1961).
- [3] D. J. BenDaniel and C. B. Duke, Phys. Rev. **152**, 683 (1966).
- [4] T. Ando and S. Mori, Surf. Sci. **113**, 124 (1982).
- [5] Q.-G. Zhu and H. Kroemer, Phys. Rev. B **27**, 3519 (1983).
- [6] The form of the effective Hamiltonian cannot be completely determined at the mesoscopic level. One requires a realistic microscopic model from which the mesoscopic approximation

may be derived. Nevertheless, the commonly used forms, obtained by placing material-dependent quantities in the center of the derivatives or taking average values for inter-atomic-plane matrix elements, appear to be adequate for many systems.

- [7] G. H. Wannier, Phys. Rev. **52**, 191 (1937).
- [8] J. C. Slater, Phys. Rev. **76**, 1592 (1949).
- [9] By “apparently hermitian,” I mean an operator in which the z -dependent parameters appear symmetrically with respect to the derivatives. The rigorous demonstration of hermiticity requires a specification of the boundary conditions applied to the operator. For an extensive discussion of the physical consequences of boundary conditions, see [20].
- [10] J. M. Luttinger and W. Kohn, Phys. Rev. **97**, 869 (1955).
- [11] G. Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures* (Halsted Press, New York, 1989), ch. 3.
- [12] M. G. Burt, Semicond. Sci. Technol. **2**, 460 (1987).
- [13] Bastard’s Hamiltonian [equation (20) of chapter 3] contains a term of the form $i\langle l|p_z|m\rangle\partial/\partial z$, which is hermitian only if this matrix element is independent of z . Burt [12] has pointed out that a rigorous derivation of the envelope-function approach requires that the same Bloch functions be used throughout the heterostructure, which would assure the hermiticity. In practice this condition is often violated as the Bloch functions are presumed to vary with the local material composition.
- [14] M. Altarelli, Phys. Rev. B **28**, 842 (1983).
- [15] J. C. Slater and G. F. Koster, Phys. Rev. **94**, 1498 (1954).
- [16] J. N. Schulman and T. C. McGill, Phys. Rev. B **19**, 6341 (1979).
- [17] D. Z.-Y. Ting, E. T. Yu, and T. C. McGill, Phys. Rev. B **45**, 3583 (1992).
- [18] A similar concept and notation for the current density appears in finite-difference numerical computations. See, for example, S. Selberherr, *Analysis and Simulation of Semiconductor Devices* (Springer, Vienna, 1984), ch. 6.
- [19] D. Z.-Y. Ting and Y.-C. Chang, Phys. Rev. B **36**, 4359 (1987).
- [20] W. R. Frensley, Rev. Mod. Phys. **62**, 745 (1990).

IV. 2015 ADDENDUM

A. Irregular Triangular Meshes

The finite-element technique has become a popular way of solving problems that are described by partial differential equations. One of the chief characteristics of this technique is the use of geometrically and topologically irregular meshes. Discussions of this approach always cast it as merely an approximation to the continuum problem, but a more sophisticated approach will seek to determine the relations that a discrete model exactly satisfies. In particular, we will ask what is the form of the continuity equation implied by the discretization of Schroedinger's equation on such an irregular mesh?

Let us assume that Schroedinger's equation has been discretized on a two-dimensional irregular triangular mesh using the "cotangent formula" for the Laplacian [21].

$$(L\psi)_i = \frac{1}{A_i} \sum_j \omega_{i,j} (\psi_j - \psi_i), \quad (18)$$

where $\omega_{i,j}$ is the sum of two cotangents of angles contained in the two triangles which share the i - j edge, and zero if the points i and j are more remotely located. Also, $\omega_{i,j} = \omega_{j,i}$ and A_i is the area of the cell enclosing point i . If we use this Laplacian in the Schroedinger equation, and construct the continuity equation in the usual way, we find:

$$\begin{aligned} \frac{\partial \rho_i}{\partial t} &\equiv \frac{\partial}{\partial t} (A_i \psi_i^* \psi_i) = -\frac{i\hbar}{2m} \sum_j [\psi_i \omega_{i,j} (\psi_j^* - \psi_i^*) - \psi_i^* \omega_{i,j} (\psi_j - \psi_i)], \\ &= -\frac{i\hbar}{2m} \sum_j \omega_{i,j} (\psi_j^* \psi_i - \psi_i^* \psi_j). \end{aligned} \quad (19)$$

Thus, the finite-element case has the same structure as the Wannier-Orbital case, with a different current density associated with each pair of coupled mesh points. Observe that we have to take into account the variation of area associated with each mesh point, to express the continuity relation in terms of the total probability within each cell.

One potential complication with this formulation is that the coupling elements $\omega_{i,j}$ are not guaranteed to be non-negative, and negative values are known to occur if oblique triangles are present in the mesh. This will produce currents that flow in the "wrong" direction, but such effects also occur in solutions of the classical diffusion equation[22].

B. Relation Between Current Density and the Velocity Operator

Another procedure which has been used to derive the current density is to identify it with the velocity operator, as defined by the Heisenberg equation of motion:

$$J \leftrightarrow v = (i/\hbar) [H, z]. \quad (20)$$

For a continuum model, this will produce the same results as (6) and (9) after symmetrization of v with respect to its adjoint. (That is, transformation to an anti-commutator superoperator.)

For discrete formulations, this procedure produces a subtle discrepancy with respect to the discrete continuity equation. Consider the simple discretization of the effective-mass Hamiltonian and position operator on a uniform mesh of spacing a :

$$H_{i,j} = \frac{\hbar^2}{2m^*a^2} (-\delta_{i-1,j} + 2\delta_{i,j} - \delta_{i+1,j}),$$

$$z_{i,j} = ja\delta_{i,j}.$$

Then,

$$v_{i,j} = -\frac{i\hbar}{m^*} \left(\frac{\delta_{i+1,j} - \delta_{i-1,j}}{2a} \right). \quad (21)$$

Thus, the use of the Heisenberg equation leads to a velocity whose values are coincident with the mesh points and which is defined by a centered difference. This is also what would be obtained by application of (20) followed by discretization according to textbook recommendations.

Applying the adjoint symmetrization, we find the explicit form:

$$\langle v \rangle_i = -\frac{i\hbar}{4m^*a} (\psi_{i+1}^* \psi_i - \psi_i^* \psi_{i+1} + \psi_i^* \psi_{i-1} - \psi_{i-1}^* \psi_i). \quad (22)$$

The tight-binding approach described above is equally applicable to this case, but it produces a current density value which is associated with the interval between meshpoints (as the $i + 1/2$ index implies), yielding the explicit expression:

$$\langle J \rangle_{i+1/2} = \frac{\hbar}{2im^*a} (\psi_{i+1}^* \psi_i - \psi_i^* \psi_{i+1}), \quad (23)$$

which will exactly satisfy the continuity equation (14). These differing results are related by:

$$\langle v \rangle_i = \frac{1}{2} (\langle J \rangle_{i-1/2} + \langle J \rangle_{i+1/2}). \quad (24)$$

In a steady-state problem in one dimension, all of the currents discussed above will be equal, and equal to those at any other location. In transient situations, however, the central-difference will not exactly satisfy any simple continuity equation. Maintaining the central-difference assumption, the net inflow of $\langle v \rangle$ will be:

$$\begin{aligned}
-\nabla \cdot \langle v \rangle_i &\leftrightarrow \frac{1}{2} (\langle v \rangle_{i-1} - \langle v \rangle_{i+1}), \\
&= \frac{1}{4} (\langle J \rangle_{i-3/2} + \langle J \rangle_{i-1/2} - \langle J \rangle_{i+1/2} - \langle J \rangle_{i+3/2}), \\
&= \frac{1}{2} \frac{\partial \rho_i}{\partial t} + \frac{1}{4} (\langle J \rangle_{i-3/2} - \langle J \rangle_{i+3/2}).
\end{aligned} \tag{25}$$

This is clearly not going to lead to any sensible continuity equation, as it invokes more remote current components. Consequently we must conclude that the Heisenberg velocity cannot be identified with the current density in discrete domains.

The expected rebuttal to the argument that I have just made is that the discrepancies will disappear as we let the mesh spacing approach zero. That cannot be taken for granted when the central-difference approximation to the gradient is employed. Because the central difference produces an anomalous value of zero for the maximum spatial Fourier component $k = \pi/a$, the convergence to the continuum is not uniform. This is the origin of the long-standing problem of anomalous states in envelope-function models of quantum states in heterostructures, and the remedy to this problem is the use of first-order differences [23].

ADDITIONAL REFERENCES

- [21] M. Meyer, *et al.*, “Discrete Differential-Geometry Operators for Triangulated 2-Manifolds,” *Visualization and Mathematics III*, Hege and Polthier, eds., pp. 35-57 (Springer, 2002).
- [22] M. Wardetzky, *et al.*, “Discrete Laplace Operators: No Free Lunch,” *Proceedings of the Fifth Eurographics Symposium on Geometry Processing*, Barcelona, Spain, 2007, pp. 33-7.
- [23] W. R. Frensky and R. Mir, arXiv 1412.7201 (2015).